**Python\_Lesson6: Python Programming**

Please don't forget to submit your feedback after the class. This helps a lot in increasing effectiveness of the course. Use the following link to submit your feedback: <https://docs.google.com/forms/d/1AkrJ-_5ZHozJf_mCg92hsXAEMYKzo3CaM4ebm0GJSmU/edit?ts=5b74fe0e>

**Lesson Overview:**

In this lesson, we will focus on clustering techniques

Clustering using KMeans and Dimension reduction using PCA

**Use Case Description:**

1. KMeans Clustering and PCA

**Programming elements:**

KMeans Clustering and Data Analysis

**Source Code:**

<https://umkc.box.com/s/pgqpapgr12wyuj63sgs3162wztpe3q77>

**In class programming:**

1. Apply K means clustering in this data set provided below:

<https://umkc.box.com/s/a9lzu9qoqfkbhjwk5nz9m6dyybhl1wqy>

* Remove any null values by the mean.
* Use the elbow method to find a good number of clusters with the KMeans algorithm

1. Calculate the silhouette score for the above clustering
2. Try feature scaling to see if it will improve the Silhouette score
3. Apply PCA on the same dataset.

Data Description can be found in <https://umkc.box.com/s/okcxw829g3m5efkljbtwpoft23bmvvnw>

**\*\*\* Bonus points**

1. **Apply kmeans algorithm on the PCA result and report your observation if the score improved or not?**
   1. **You can try different variation like PCA+KMEANS , SCALING+PCA+KMEANS.**
2. **Visualize the clustering of first bonus question**

**ICP Submission Guidelines**

**(for In Class students):**

1. ICP Submission is in pairs of two students.

2. Once completed, must be presented to TA or Instructor before the completion of the class

3. Submission after class is considered as a late submission. (Check the late submission policy in the syllabus)

4. ICP Code with brief explanation should be pushed to GitHub. Submit GitHub link through the Feedback Form: <https://goo.gl/forms/HEJyYaiUi3MKXKP22>

**Online Submission Guidelines (for Online students):**

1. Submit your source code and documentation to GitHub and represent the work through wiki page properly (submit your screenshots as well. The screenshot should have both the code and the output)

2. Comment your code appropriately

3. Video Submission (2 – 3 min video showing the demo of the ICP, with a brief voice over on the code explanation)

4. Submission after class is considered as a late submission. (Check the late submission policy in the syllabus)

5. Use the following Google link to submit your ICP # (GitHub wiki page link for ICP #): https://docs.google.com/forms/d/1AkrJ-\_5ZHozJf\_mCg92hsXAEMYKzo3CaM4ebm0GJSmU/edit?ts=5b74fe0e

**Evaluation Criteria:**

1. Completeness of Features

2. Code Quality (<https://en.wikipedia.org/wiki/Best_coding_practices>)

3. Time

4. Feedback Submission

**Note:** *Cheating, plagiarism, disruptive behavior and other forms of unacceptable conduct are subject to strong sanctions in accordance with university policy. See detailed description of university policy at the following URL:* [*https://catalog.umkc.edu/special-notices/academic-honesty/*](https://catalog.umkc.edu/special-notices/academic-honesty/)